

能動学習と Few-Shot 学習を用いた短文自動採点システムのための微調整手法の提案

堀 達也、山内 康一郎 (中部大学大学院)

1 背景・目的

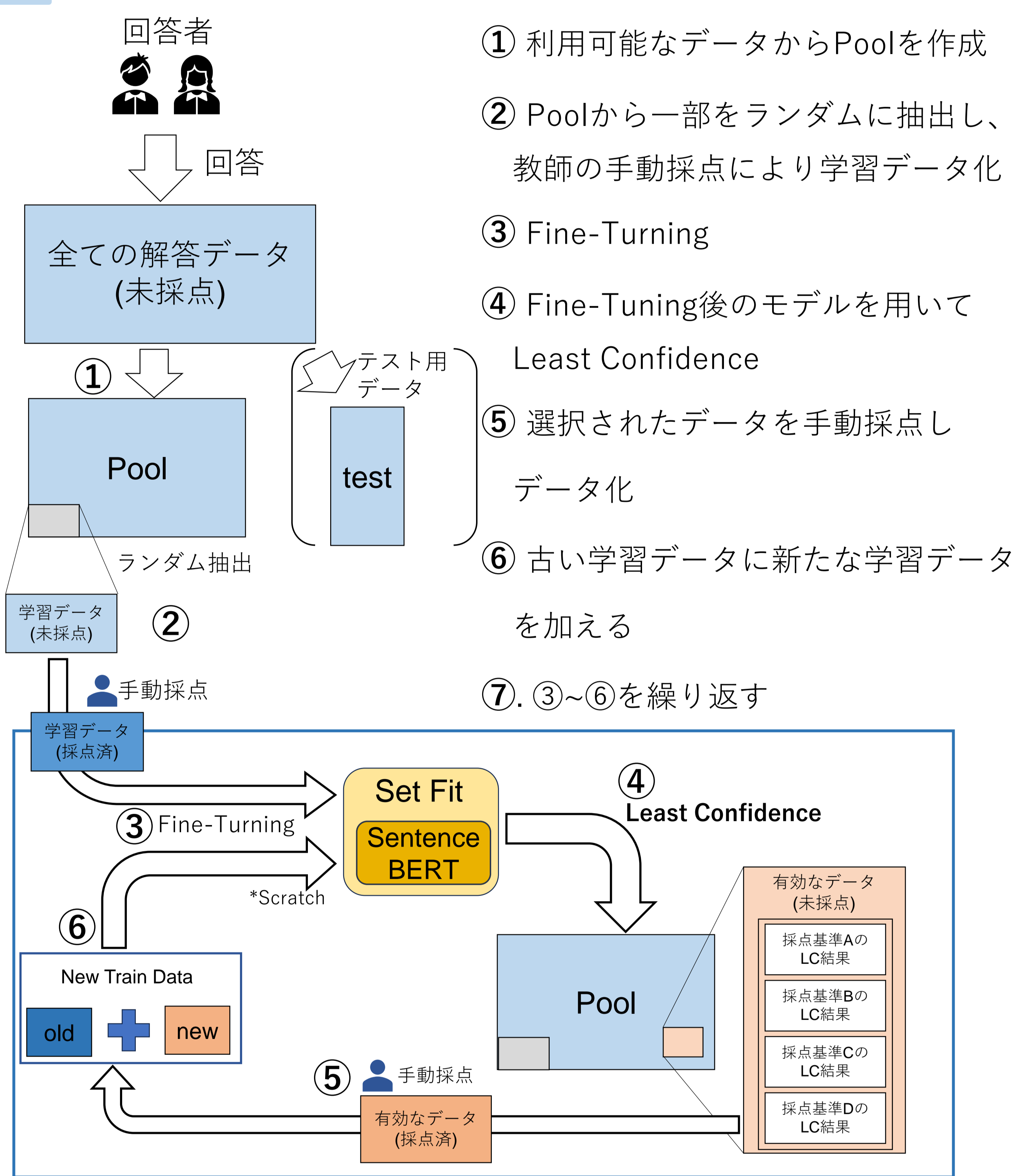
短文記述問題とは：回答者が短文を作成して回答する問題形式
短文記述問題の自動採点の課題

- 課題1：学習データ作成のためのアノテーション(手動採点)に大きな労力がかかる
- 課題2：学校の教員では大量の採点データを集めることが困難
 ⇒学校の教員によるテストは学年や科目で細分化されている(日本の高校における高校3年生の平均在籍数：約225人)

研究目的

- 目的1**：Few-Shot学習を用いた自動採点の実現
- 目的2**：現実の採点状況(利用可能なデータが少ない)における自動採点の検証

2 提案手法



■ Set Fit^[1]

Sentence BERT^[2]のFine turning 時のFew-Shot学習用フレームワーク。過学習を起こしやすい。

■ Least Confidence (LC)

- 能動学習手段の一種。
- Set Fitの過学習を防ぎ、効率的な学習を行う狙い。

$$NewData = \operatorname{argmin}_x (\max_y P(y|x))$$

3 利用データセット

- 理研記述問題データセット^[3](2100件)を利用
- 回答データと各採点基準における採点結果が含まれる

データ例(使用データ：Y14_1-2_2_4)

回答データ	採点基準				合計点
	A	B	C	D	
鰻屋の座敷の空気に馴染み、溶けこみきって、「鰻屋の風景」になってい...	0点 (0/3)	0点 (0/2)	3点 (3/4)	3点 (3/3)	6点 (6/10)

4 実験設定

性能検証

LC：提案フレームワークを用いて新規学習データを追加
 Without LC：新規学習データをランダムに抽出して追加

2種類の実験を実施

- 実験1：Poolデータ1600件,テストデータ500件
 ⇒ Few-Shot学習を用いた自動採点の検証(**目的1**)
- 実験2：Poolデータ200件の中から学習&テスト
 ⇒現実的な状況における検証(**目的2**)

評価指標：Accuracy, RMSE

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

y_i : 真の値, \hat{y}_i : 推定した値

使用モデル：公開されているSentence BERT^[4]

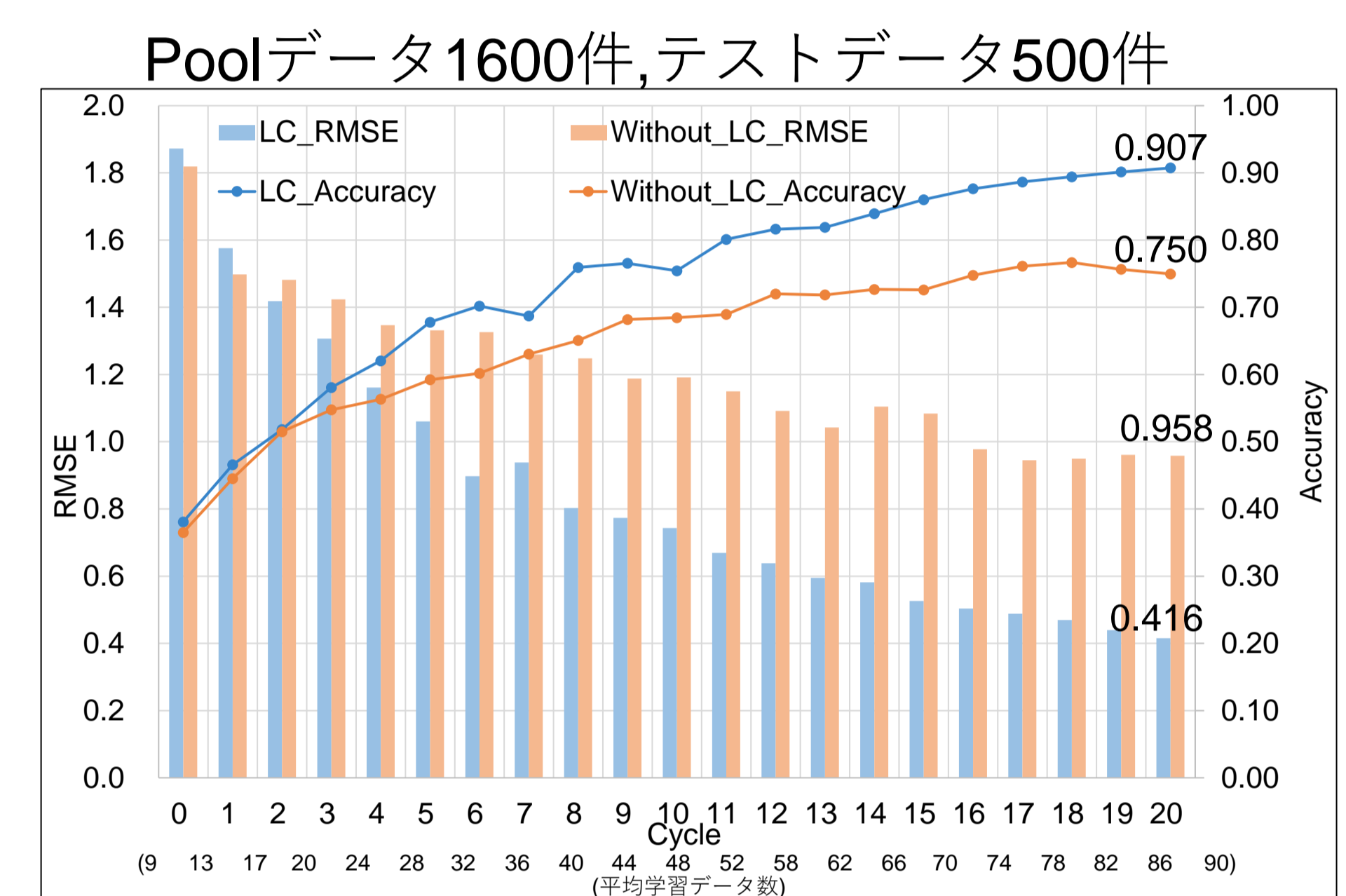
5 実験結果

実験1

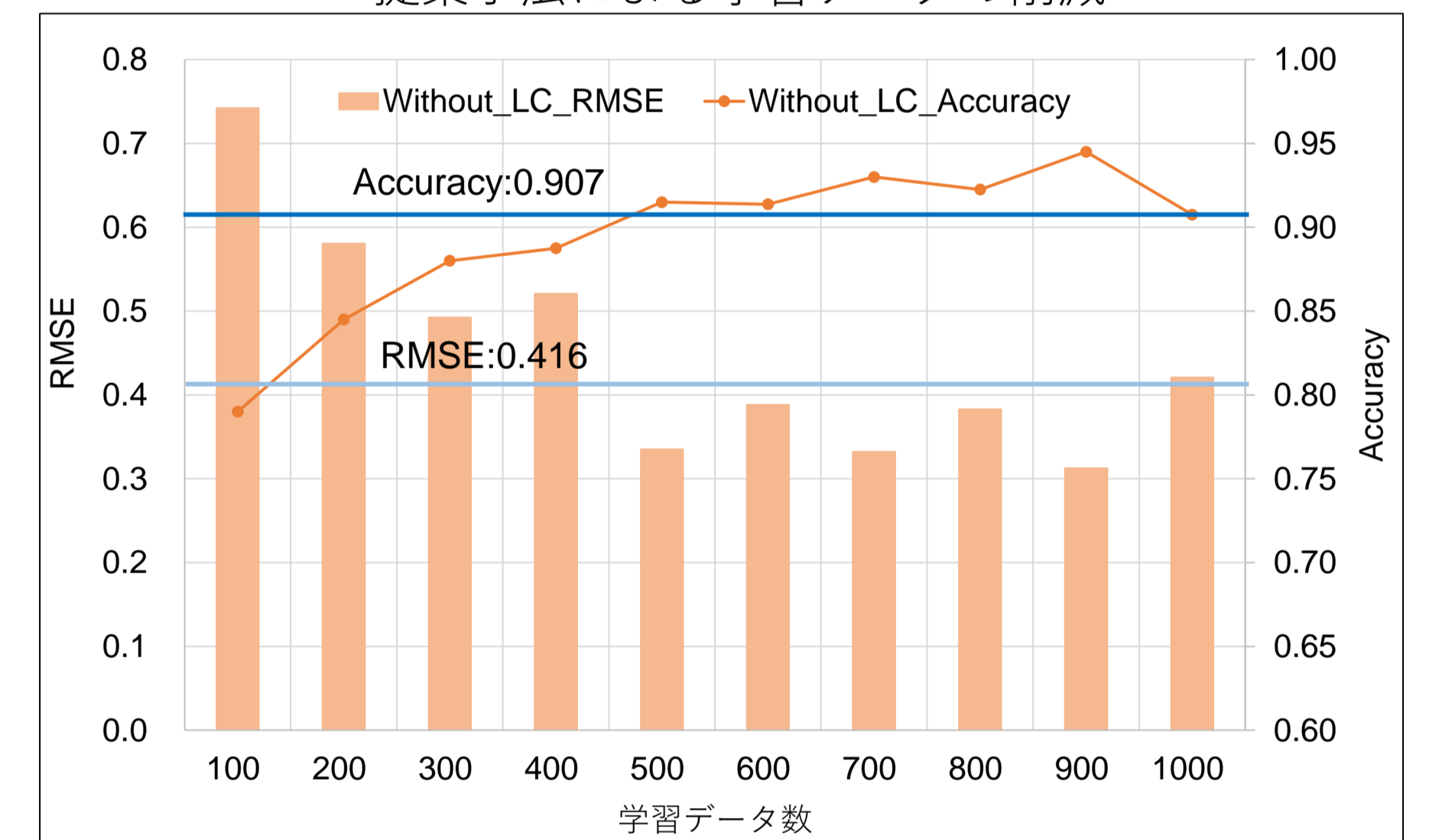
Accuracy
 Without LC : 0.750
 LC : 0.907

RMSE
 Without LC : 0.958
 LC : 0.416

LC(90件)はWithout LC
 における400~500
 データ相当 (80.0%削減)



提案手法による学習データの削減

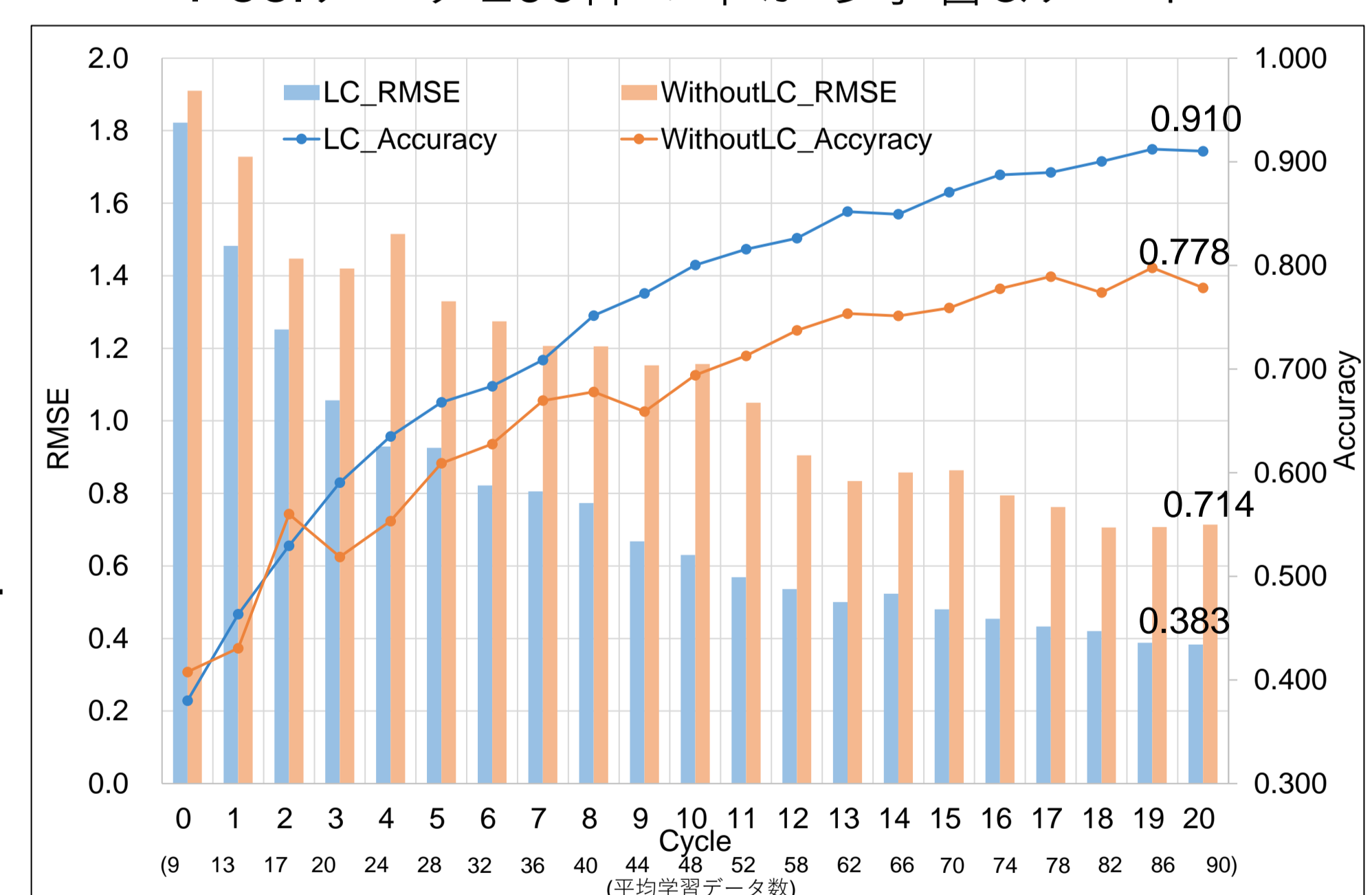


実験2

Accuracy
 Without LC : 0.778
 LC : 0.910

RMSE
 Without LC : 0.714
 LC : 0.383

Poolデータ200件の中から学習&テスト



6 まとめ今後の課題

まとめ

- Few-Shot学習と能動学習を組み合わせた自動採点用のフレームワークを作成した。
- 提案手法により、約80%の学習データ削減を達成した。
- データ数制限下では、非制限下と比較してわずかな性能が低下となった。

課題

- 計算量の削減⇒使用モデルや追加学習手法の検討
- 実用化への課題⇒必要精度の検討

参考文献

[1] Lewis Tunstall, Efficient Few-Shot Learning Without Prompts, <https://doi.org/10.48550/arXiv.2209.11055>
 [3] NII理研記述問題データセット, <https://www.nii.ac.jp/dsc/idr/rdata/RIKEN-SAA/>

[2] Nils Reimers, Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, <https://doi.org/10.48550/arXiv.1908.10084>
 [4] sentence-transformers/paraphrase-mpnet-base-v2, <https://huggingface.co/sentence-transformers/paraphrase-mpnet-base-v2>